

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re U.S. Patent Application of

KITAHARA *et al.*

Application Number: To be assigned

Filed: Concurrently herewith

For: METHOD FOR SPEECH INTERPRETATION
AND SPEECH INTERPRETATION SERVER

Honorable Assistant Commissioner
for Patents
Washington, D.C. 20231

Priority

1c971 U.S. PTO
09/811442



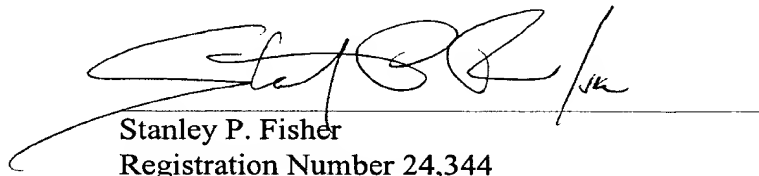
**REQUEST FOR PRIORITY
UNDER 35 U.S.C. § 119
AND THE INTERNATIONAL CONVENTION**

Sir:

In the matter of the above-captioned application for a United States patent, notice is hereby given that the Applicant claims the priority date of October 17, 2000, the filing date of the corresponding Japanese patent application 2000-321921.

The certified copy of corresponding Japanese patent application 2000-321921 is being submitted concurrently herewith. Acknowledgment of receipt of the certified copy is respectfully requested in due course.

Respectfully submitted,


Stanley P. Fisher
Registration Number 24,344

REED SMITH HAZEL & THOMAS LLP
3110 Fairview Park Drive
Suite 1400
Falls Church, Virginia 22042
(703) 641-4200

JUAN CARLOS A. MARQUEZ
Registration No. 34,072

March 20, 2001

310001547

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

JC971 U.S. PTO
09/811442
03/20/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
in this Office.

出願年月日
Date of Application:

2000年10月17日

CERTIFIED COPY OF
PRIORITY DOCUMENT

願番号
Application Number:

特願2000-321921

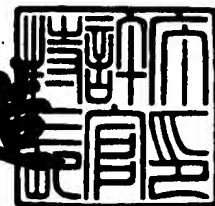
願人
Applicant(s):

株式会社日立製作所

2001年 2月23日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2001-3010329

2001-3010329

【書類名】 特許願

【整理番号】 H00015471A

【提出日】 平成12年10月17日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 17/28

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

【氏名】 北原 義典

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

【氏名】 大淵 康成

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

【氏名】 小泉 敦子

【発明者】

【住所又は居所】 神奈川県川崎市幸区鹿島田890番地
株式会社日立製作所コンシューマネットビジネス推進本部内

【氏名】 水谷 世希

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【電話番号】 03-3212-1111

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声通訳サービス方法および音声通訳サーバ

【特許請求の範囲】

【請求項 1】

通訳サーバとモバイルインターネットアクセス網で接続された携帯端末の画面に該通訳サーバに格納された通訳用文例集を表示させるステップと、
前記携帯端末から音声通訳サービスへの接続指示を受け前記携帯端末と前記通訳サーバとの接続を電話網接続に切り換えるステップと、
前記携帯端末に入力されたユーザ音声を電話網を通じて受けとるステップと、
前記ユーザ音声を認識するステップと、
認識したユーザ音声を異言語表現に通訳するステップと、
通訳した異言語表現に対応する音声を電話網を通じて前記携帯端末に出力するステップとを有することを特徴とする音声通訳サービス方法。

【請求項 2】

請求項 1 に記載の音声通訳サービス方法において、
さらに、前記モバイルインターネットアクセス網を通じて、前記ユーザが通訳可能な言語種を、前記携帯端末の表示画面にメニューとして表示させるステップと、
ユーザから該言語種メニューから通訳先言語の選択を受けるステップと、
認識したユーザ音声を選択された言語の表現に通訳するステップとを有することを特徴とする音声通訳サービス方法。

【請求項 3】

請求項 1 に記載の音声通訳サービス方法において、
さらに、前記モバイルインターネットアクセス網を通じて通訳可能な文例分野を前記携帯端末の表示画面に表示させるステップと、
ユーザから該分野メニューから通訳可能文例場面の選択を受けるステップと、
認識したユーザ音声を選択された文例場面に即した言語表現に通訳するステップとを有することを特徴とする音声通訳サービス方法。

【請求項 4】

請求項 1 に記載の音声通訳サービス方法において、

ユーザ音声認識の結果を電話網を通じてユーザの携帯端末に音声出力するステップを有することを特徴とする音声通訳サービス方法。

【請求項 5】

請求項 4 に記載の音声通訳サービス方法において、
前記認識した結果を携帯端末で受けたユーザからの指示を受けた後、通訳音声を電話網を通じて前記携帯端末に出力することを特徴とする音声通訳サービス方法。

【請求項 6】

請求項 5 に記載の音声通訳サービス方法において、前記指示は、前記携帯端末のボタンを押し下げることによって行われ、前記ユーザが、前記ボタンを押下する度に、通訳音声を出力することを特徴とする音声通訳サービス方法。

【請求項 7】

請求項 5 に記載の音声通訳サービス方法において、前記指示は、前記ユーザによる特定の単語もしくは文音声の発声入力により行われ、前記指示により通訳音声を出力することを特徴とする音声通訳サービス方法。

【請求項 8】

請求項 7 に記載の音声通訳サービス方法において、
前記ユーザによるある特定の単語もしくは文音声を発声、入力を受ける度に、通訳音声を出力することを特徴とする通訳サービス方法。

【請求項 9】

請求項 1 に記載の音声通訳サービス方法において、
音声認識に用いる辞書もしくは文法もしくはこれらの両方を前記携帯端末の表示画面に表示された文例に関する項目に絞り込んで音声認識を行なうことを特徴とする音声通訳サービス方法。

【請求項 10】

請求項 1 から請求項 5 のいずれかに記載の自動音声通訳サービス方法において、
前記通訳音声を前記携帯端末に出力する回数に無関係に、ユーザに課金することを特徴とする音声通訳サービス方法。

【請求項 11】

電話網を通じて音声の入力を受ける音声入力手段と、
 入力された音声認識し所定の記号列に変換する音声認識手段と、
 変換された記号列を前記入力された音声の言語と異なる言語の表現に変換する言語変換手段と、
 変換された言語表現に対応する音声を出力する音声出力手段と、
 通訳用文例を格納する文例テーブルと、
 モバイルインターネットアクセス網で接続された携帯端末の画面に前記通訳用文例を表示させる文例表示手段とを有することを特徴とする音声通訳サーバ。

【請求項 12】

前記音声認識手段は、前記文例テーブルを参照して通訳用の文例と、動作を指示するコマンド文とを区別して認識することを特徴とする音声通訳サーバ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ユーザが発声した音声进行他の言語に翻訳し音声として出力する自動通訳サービスに係るものである。主としてモバイルインターネットアクセスサービス利用可能な携帯電話を用いたサービスを対象とするが、通常の有線電話を用いたり、パーソナルコンピュータ、携帯情報端末等を用いて前記動作を行なう音声通訳サービスも含む。

【0002】

【従来の技術】

従来、音声通訳機というものがあつた。例えば、特開平11-125539には、内蔵マイクロホンより入力されて音声を異なる言語に翻訳し、翻訳結果をマイクロホンとは反対の面に設置されたスピーカより音声出力する、小型かつ片手操作可能な音声通訳装置が開示されている。しかしながら、このような音声通訳装置は専用機のため、他の用途には利用できず、海外旅行などの实际利用場面では、結局ユーザの所持品を増やしてしまうことになる。一方、例えば、特開平9-65424には、音声認識サーバと無線携帯端末を用いた音声翻訳システムが

開示されている。しかしながら、この音声翻訳システムでは、任意文を入力対象としているため、高い通訳精度が得られにくく、実際の運用が困難である、という問題があった。

【0003】

【発明が解決しようとする課題】

本発明の目的は、通話用の電話機、とりわけ、普及率の高いモバイルインターネットアクセスサービス利用可能な電話を用いて、精度の高い音声通訳が誰でも簡単に利用できるサービスを提供することにある。

【0004】

【課題を解決するための手段】

本発明では、ユーザが電話を用いてサーバ上で稼動する自動通訳サーバに音声を送信すると、翻訳された音声、再びユーザの電話に返信されるようにする。その際に、最初にユーザがモバイルインターネットアクセスサービス利用可能な電話からモバイルインターネットアクセスサービスパケット網を通じて、モバイルインターネットアクセスサービスゲートウェイサーバに接続すると、前記自動通訳サーバが、ユーザが通訳可能な言語種を、該ユーザの電話の表示画面にメニューとして表示し、ユーザが該言語種メニューから通訳先言語を選択できるようにする。また、ユーザが通訳可能な文例場面を該ユーザの電話の表示画面に表示し、ユーザが該場面メニューから通訳可能文例場面を選択できるようにする。また、ユーザが入力可能な文例を該ユーザの電話の表示画面に表示し、ユーザが該入力可能な文例を見ながら発声入力できるようにする。さらに、前記自動通訳サーバは、該入力された音声を前記表示した範囲の文例辞書を用いて認識し、訳語に変換し、音声として前記電話端末に出力する。

【0005】

このように、新たに専用機を購入することなく一般の電話端末を用い、該電話端末に表示した定型的な文を入力対象として、自動通訳サーバで通訳処理をすることによって、通訳精度が極めて高い通訳サービスが容易に受けられる環境を提供することができる。

【0006】

【発明の実施の形態】

以下、本発明の実施例を図を用いて説明する。

【0007】

図1は本発明の第1の実施例を示す自動通訳サービスシステム構成図である。本発明は、音声通訳機に関するものであり、本実施例におけるサーバは、パソコン、ワークステーション等、中央演算装置、メモリを備え、同図のような構成をなすことができるものであれば何でもよく、プラットフォームの種類が本発明の適用範囲を限定するものではない。また、本実施例は英語を日本語に通訳する自動通訳サービスであるが、これは一例であり、日英、中日、日中等、言語の種類は限定されない。

【0008】

同図において、1はモバイルインターネットアクセスサービス利用可能な電話端末、2はモバイルインターネットアクセスサービスパケット網、3はモバイルインターネットアクセスサービスゲートウェイサーバ、4は電話回線ボード、5は音声入力手段、6は音声認識手段、7は言語翻訳手段、8は単語辞書、9は文法テーブル、10は言語変換テーブル、11は音声生成手段、12は素片セット、13は音声出力手段、14は中央演算装置、15はメモリ、16は言語種表示手段、17は場面表示手段、18は文例表示手段、19は認識候補表示手段、20は文辞書、21はコマンド辞書、22は言語種テーブル、23は場面テーブル、24は文例テーブル、31は認証サーバ、32は課金サーバ、33は課金テーブルである。メモリ15のデータ構造を図2に示す。また、モバイルインターネットアクセスサービス利用可能な電話端末1の概観の例を図3に示す。なお、ここでは、モバイルインターネットアクセスサービス利用可能な電話端末として、まず、通話音声およびデータを同一のプロトコルで扱える電話端末を仮定し説明する。

【0009】

図1、図3において、まず、モバイルインターネットアクセスサービス利用可能な電話端末1の電源104をいれ、モバイルインターネット接続用ボタン102を押下すると、モバイルインターネットアクセスサービスパケット網2を介して、センタの

モバイルインターネットアクセスサービスゲートウェイサーバ3に接続され、認証サーバ31でモバイルインターネットアクセスサービス利用登録ユーザであることが確認される。この時点で、ユーザIDが課金サーバ32に送られる。当然のことながら、ユーザIDは電話端末1のIDとリンクしていることが多いが、電話端末1とは無関係にユーザIDをパスワードと組み合わせて入力させる方法もある。

【0010】

課金サーバ32は課金テーブル33をもつ。課金テーブル33のデータ構造を図11に示す。認証サーバ31から送られたIDは、課金テーブル33のユーザIDカラム411の各セルと照合され、一致したセルの料金カラム412がゼロリセットされる。例えば前記ユーザのユーザIDが"1236"であるとする、課金テーブル33のユーザIDカラム411のセルのうち"1236"であるセル403と一致するので、料金カラム412の403に対応するセルが"0"となる。

【0011】

モバイルインターネットアクセスサービスゲートウェイサーバ3には、本実施例の自動通訳サーバのような各種サービスサーバが専用線で接続されていたり、一般のインターネット網が接続されている。なお、モバイルインターネットアクセスサービスゲートウェイサーバ3への接続には、前記のようなモバイルインターネット接続用ボタン102を押下する方法のほかに、ある所定の電話番号を数字ボタンで入力する方法等を用いてもよい。

【0012】

モバイルインターネットアクセスサービス利用可能な電話端末1が、モバイルインターネットアクセスサービスゲートウェイサーバ3に接続され、認証サーバ31でモバイルインターネットアクセスサービス利用登録ユーザであることが確認されると、モバイルインターネットアクセスサービスゲートウェイサーバ3は、電話端末1のディスプレイ101に、図4のようにサービスのメニューを表示する。該サービスメニューは、図4のように、白黒反転により1番目の項目の輝度が他の項目とは異なる状態で初期表示される。これは、第1番目の項目が選択されていることを示すが、もちろん、その表示は白黒反転に限らず、他の項目と区別がつけられるような方法であればよい。

【 0 0 1 3 】

ユーザは、該サービスメニューを見ながら電話端末1上の上下カーソル移動の機能を割り当てた所定のボタンを押して、白黒反転を第3番目の項目”自動通訳”まで移動させて、さらに電話端末1上の前記所定のボタンとは別の確定の機能を割り当てた所定のボタンを押下し確定する。メニュー項目”自動通訳”が確定すると、電話端末1がモバイルインターネットアクセスサービスゲートウェイサーバ3経由で自動通訳サーバ1000に接続される。

【 0 0 1 4 】

次に、自動通訳サーバ1000の言語種表示手段16が起動され、言語種テーブル22を用いて電話端末1上のディスプレイ101に、図5のように通訳可能な言語の種類を表示する。言語種テーブル22は図12のようなデータ構造をもっており、言語種表示手段16は該テーブルの言語種812の各項目を、前記電話端末1に送り図5のように該電話端末1上のディスプレイ101に表示する。同図では、白黒反転により1番目の項目の輝度が他の項目とは異なる状態で初期表示されている。これは、第1番目の項目が選択されていることを示すが、もちろん、その表示は白黒反転に限らず、他の項目と区別がつけられるような方法であればよい。ユーザは、該言語種メニューを見ながら電話端末1上の上下カーソル移動の機能を割り当てた所定のボタンを押して、白黒反転を例えば第2番目の項目”日本語→英語”まで移動させて、さらに電話端末1上の前記所定のボタンとは別の所定の確定の機能を割り当てたボタンを押下し確定する。この時、言語種表示手段16は電話端末1のカーソルの位置を受け取りメモリ15上のLANG209に該位置番号を格納する。例えば、日本語を英語に通訳したい場合には、「日本語→英語」が2行目にあるので、メモリ15上のLANG209に”2”が格納される。

【 0 0 1 5 】

もちろん、この言語種指定は、電話端末1上のディスプレイ101に言語種を表示し、ユーザに所望の言語種を上下移動ボタンで選択させる方法の他に、前記電話端末1上のディスプレイ101をタッチパネルディスプレイで構成し指等の接触により所望の言語種を指定する方法を用いるようにしてもよい。

【0016】

また、あらかじめある所定の電話番号を言語種毎に割り当てておき、ユーザが所望の言語種に対応する電話番号を前記電話端末1上の数字ボタンで入力する方法等を用いてもよい。

【0017】

言語種が確定すると、次に、自動通訳サーバ1000の場面表示手段17が起動され、場面テーブル23を用いて、電話端末1上のディスプレイ101に、図6のように通訳可能な場面を表示する。ここでいう場面とは、「空港」「ホテル」「レストラン」等、本発明の通訳サービスを利用する場面のことである。場面テーブル23は、図13のようなデータ構造をもっており、場面表示手段17は該テーブルの場面912の各項目を、前記電話端末1に送り図6のように該電話端末1上のディスプレイ101に表示する。同図では、白黒反転により1番目の項目の輝度が他の項目とは異なる状態で初期表示されている。これは、第1番目の項目が選択されていることを示すが、もちろん、その表示は白黒反転に限らず、他の項目と区別がつけられるような方法であればよい。

【0018】

ユーザは、該場面メニューを見ながら電話端末1上の上下カーソル移動の機能を割り当てた所定のボタンを押して、白黒反転を例えば第3番目の項目”レストラン”まで移動させて、さらに電話端末1上の前記所定のボタンとは別の確定の機能を割り当てた所定のボタンを押下し確定する。この時、場面表示手段17は電話端末1のカーソルの位置を受け取りメモリ15上のSCENE210に該位置番号を格納する。例えば、レストランの場面で通訳したい場合には、「レストラン」が3行目にあるので、メモリ15上のSCENE210に”3”が格納される。

【0019】

もちろん、この場面指定は、電話端末1上のディスプレイ101に場面を表示し、ユーザに所望の場面を上下移動ボタンで選択させる方法の他に、前記電話端末1上のディスプレイ101をタッチパネルディスプレイで構成し指等の接触により所望の場面を指定する方法を用いるようにしてもよい。また、あらかじめある所定の電話番号を場面毎に割り当てておき、ユーザが所望の場面に対応する電話番号

を前記電話端末1上の数字ボタンで入力する方法等を用いてもよい。

【 0 0 2 0 】

場面が確定すると、次に、自動通訳サーバ1000の文例表示手段18が起動され、文例テーブル24を用いて電話端末1上のディスプレイ101に、図7のように通訳可能な文例を表示すると同時に、自動通訳サーバ1000の音声入力手段5が起動される。続いて、音声入力手段5は、システムを音声入力可能な状態にする。その後、ユーザは、該文例を見ながら、例えば先の例であれば、日本語でレストランの場面での通訳させたい文を電話端末1の送話口のマイク104に向かって発声入力する。ここでは、例として、レストランの場面で「水が欲しいです。」という文を英語に通訳したい場合を考えることにする。

【 0 0 2 1 】

文例テーブル24は図14のようなデータ構造をもっており、文例表示手段18は、文例テーブル24の場面番号511の各項目のうち、メモリ15上のSECENE210に格納されている値の項目の文例513を文例番号512の"1"から順次、前記電話端末1に送り該電話端末1上のディスプレイ101に表示する。先の例では、メモリ15上のSCENE210に"3"が格納されているので、図14の文例テーブル24の場面番号511の値が"3"である項目501,502,503,504の文例513の内容「こんにちは」「ありがとう」「[]はどこですか」「[]が欲しいです」を前記電話端末1に送り図7のように該電話端末1上のディスプレイ101にM個ずつ順次表示する。Mはディスプレイ101の大きさに応じて設定する定数であるが、ここではM=4とする。

【 0 0 2 2 】

ユーザは、前記文例の中に「[]が欲しいです」という文例があるので、この文例を基に「水が欲しいです」と発声入力する。この発声入力に際して、電話端末1上の発声入力の合図の意味をもたせた所定のボタンを押下して自動通訳サーバ1000の音声入力手段5が音声入力可能な状態にしてから発声するようにしても良いし、特にボタン等のトリガを用いなくて自動通訳サーバ1000の音声入力手段5が起動された後いつでも音声入力可能な状態にしておいても良い。なお、表示する文例は、上述の例のようにスロット[]を含有した文もしくは文法でも良いし、もしくは文そのものでも良い。ここで、スロットとは、単語、句などをその個所

に設定できるといういわばボックスである。例えば、「[]が欲しいです」では、スロット[]に「水」「コーヒー」「冷たい水」等を設定できる。

【 0 0 2 3 】

このように、文例を表示し文型をある程度規定することにより、音声認識の精度を著しく向上させることができる。また、文例を表示することでユーザの使い勝手を向上させることができる。

【 0 0 2 4 】

なお、前記表示文例は、上下カーソル移動の機能を割り当てた所定のボタンを押すことにより、次々とスクロールし認識可能な文例を全て順次表示させることができる。ただし、この時に、前記電話端末1上のディスプレイ101に表示させている最初の文例の文例番号512の値、および、表示させている最後の文例の文例番号512の値をメモリ15上のBSENTENCE211、ESENTENCE212に各々格納する。図7の例では、BSENTENCE211には"1"、ESENTENCE212には"4"が格納される。

【 0 0 2 5 】

続いて、音声入力手段5は、電話回線ボード4にてアナログ／デジタル変換された入力音声をメモリ15上のWAVE 201に格納する。電話回線ボード4におけるアナログ／デジタル変換のサンプリングレートは8 kHz、11 kHz、16 kHz等、ユーザが適宜定めることができる。

【 0 0 2 6 】

なお、この時点で、前記ユーザが、入力した音声を取り消し、再入力したいとするならば、電話端末1上の取り消し機能を割り当てた所定のボタンを押下することによって、音声再入力可能な状態にすることもできる。電話端末1上の取り消し機能を割り当てた所定のボタンは、押下されると、発声入力の合図の意味をもたせた所定のボタンを押下したときもしくは自動通訳サーバ1000の音声入力手段5が起動された直後と同じ初期状態にリセットされ、音声入力可能な状態になるようにする。

【 0 0 2 7 】

次に、音声認識手段6が起動される。音声認識手段6は、メモリ15上WAVE201に格納された音声データを読み込み、特徴ベクトルの時系列に変換し、各文音声の

該特徴ベクトル時系列をもつ文辞書を用いて照合を行ない、該音声データを認識し候補を出力する。特徴ベクトル時系列への変換方法や、照合方法を含めて音声認識する方法としては、例えば L. Rabiner & B.-H. Juang 著、古井貞熙 監訳「音声認識の基礎（下）」（NTTアドバンステクノロジー、1995）P 2 4 5～P 3 0 4 記載の方法などを利用することができる。もちろん、他の音声認識の方法を用いてもよく、音声認識の方法が本発明を限定するものではない。文辞書20のデータ構造を図15に示す。音声認識手段6は、メモリ15上WAVE201に格納された音声データを読み込み、文辞書20の文例番号611の値がメモリ15上のBSENTENCE211、ESENTENCE212に各々格納されている値の範囲内である項目のもの全ての特徴ベクトル時系列614の値を用いて音声認識を行なう。先の例では、BSENTENCE211には"1"、ESENTENCE212には"4"が格納されているので、文辞書20の文例番号611の値が"1"から"4"までの各項目の特徴ベクトル時系列614の値を用いて音声認識を行なう。その結果、前記音声は、尤度の高い順に、例えば、文字列"メニューがほしいです"、"みずがほしいです"、"コーヒーがほしいです"、"スプーンがほしいです"の文例番号および文番号に変換される。そのために、これらの候補の文例番号611、文番号612、および文613を尤度の高い順に、メモリ15上のRECOGNUM(1)、RECOGSNUM(1)、RECOGS(1)、RECOGNUM(2)、RECOGSNUM(2)、RECOGS(2)、…、RECOGNUM(N)、RECOGSNUM(N)、RECOGS(N)205に格納していく。ここで、Nは、文辞書20の文例番号611の値がメモリ15上のBSENTENCE211、ESENTENCE212に各々格納されている値の範囲内である項目のもの全数である。次に、認識候補表示手段19が起動され、RECOGS(1)、RECOGS(2)、…、RECOGS(M)205の内容を、図8のように、電話端末1に送り該電話端末1上のディスプレイ101に順次表示する。このとき、メモリ15上のICNT204に"1"を格納し、電話端末1上のディスプレイ101に表示されているRECOGS(ICNT)の内容が白黒反転されるようにする。Mはディスプレイ101の大きさに応じて設定する定数であるが、ここではM=4とする。また、メモリ15上のINDEX215に"0"を格納する。

【0028】

ユーザは、第1候補の表示が自分の発声した内容と一致しているもしくは近ければ、そのまま確定の機能を割り当てた所定のボタンを押下し確定する。第1候補

の表示が正しくない場合、ユーザは、上下カーソル移動の機能を割り当てた所定のボタンを押下することによって、電話端末1上のディスプレイ101上の正しい文字列が表示されている部分にまで白黒反転表示領域を下方に移動させる。すなわち、下移動ボタンを押下するたびに、メモリ15上のICNT 204の値をインクリメントし、電話端末1上のディスプレイ101上の同じくメモリ15上のRECOG(ICNT)の内容が表示されている部分のみを白黒反転表示するようにする。ここで、ICNT 204の値がMを超えた場合は、メモリ15上のINDEX215の値に”M”を加算し、次のM個の候補RECOGS(INDEX+1)、RECOGS(INDEX+2)、…、RECOGS(INDEX+M)をメモリ15上の中から読み出し、電話端末1に送り該電話端末1上のディスプレイ101に順次表示する。このとき、メモリ15上のICNT204に”1”を格納し、電話端末1上のディスプレイ101に表示されているRECOGS(INDEX+1)、RECOGS(INDEX+2)、…、RECOGS(INDEX+M)のうちのICNT番目の表示が白黒反転されるようにする。以下、同様にして、次のM個の候補を、電話端末1に送り該電話端末1上のディスプレイ101に順次表示していく。なお、図15に示した音声認識に使用する文辞書20の構造は一例であり、例えば、文法と単語辞書を組み合わせる方法もある。その場合も、メモリ15上のSCENE210に格納された場面番号に対応させて、場面別の文法や辞書を用いるようにする。

【0029】

また、上移動ボタンで白黒反転表示領域を上方に移動させるようにし、バックして表示内容を見られるようにしても良い。すなわち、上移動ボタンを押下するたびに、メモリ15上のICNT 204の値をデクリメントし、電話端末1上のディスプレイ101上のRECOGS(INDEX+1)、RECOGS(INDEX+2)、…、RECOGS(INDEX+M)のうちのICNT番目の表示部分のみを白黒反転表示するようにする。

【0030】

この候補の選択は、前記電話端末1上の上下移動ボタンを利用する方法に限らず、前記電話端末1上のディスプレイ101をタッチパネルディスプレイで構成し指等の接触により候補を指定する方法を用いるようにしてもよい。

【0031】

ユーザが、白黒反転されている表示が自分の発声した内容と一致しているもしくは

は近いとして確定の機能を割り当てた所定のボタンを押下し確定すると、メモリ15上のRECOGNUM(INDEX+ICNT)、RECOGSNUM(INDEX+ICNT)、RECOGS (INDEX+ICNT)の値を、同じくメモリ15上のPCAND213、SCAND214、JAPANESE 203に各々格納する。

【0032】

図8の例では、“みずがほしいです”が、第2行目に表示されているので、下移動ボタンを押下することにより第2行目まで、白黒反転領域を移動させ、確定ボタンを押下する。そして、INDEXは“0”ICNTは“2”となり、RECOGNUM(2)、RECOGSNUM(2)、RECOGS (2)の値である“4”、“1”、“みずがほしいです”がメモリ15上のPCAND213、SCAND214、JAPANESE 203に各々格納される。

【0033】

自分の発声した内容の確認方法には、以上述べたような電話端末1上のディスプレイ101上に音声認識候補を表示する方法の他に、ディスプレイ101を使用せずに、以下のような方法を用いることもできる。すなわち、音声認識手段6が尤度の高い順に候補の文例番号611、文番号612、および文613を尤度の高い順に、メモリ15上のRECOGNUM(1)、RECOGSNUM(1)、RECOGS(1)、RECOGNUM(2)、RECOGSNUM(2)、RECOGS(2)、…、RECOGNUM(N)、RECOGSNUM(N)、RECOGS(N)205に格納した後、音声生成手段12が起動され、このとき、メモリ15上のJCNT208に“1”を格納し、メモリ15上のRECOGS (JCNT)をまず読み出し、該文字列を合成音声に変換し、該音声の波形データをデジタル／アナログ変換によりアナログデータに変換し、音声出力手段13を通じて音声として電話端末1に送る。文字列を合成音声に変換するためには、例えば、J.Allen, M.S. Hunnicutt and D. Klatt他“From Text to Speech”(Cambridge University Press, 1987) P.16からP.150に記載されているホルマント合成方式や、矢頭「テキスト音声合成技術の最新状況」(Interface, Dec., 1996) P.161からP.165に記載されている波形重畳方式が使える。もちろん、他のテキスト音声合成方式を使うこともできる。また、認識可能な文例に対応する音声をあらかじめ録音して、メモリ等の記憶装置に格納しておくようにしてもよい。

【0034】

ユーザは、前記送られた音声を電話端末1上のスピーカ100から聞き、該音声が自分の発声入力した内容に合致していれば、電話端末1上の確定の機能を割り当てた所定のボタンを押下し確定する。該音声が自分の発声入力した内容に合致していない場合には、ユーザは電話端末1上の次候補提示の機能を割り当てた所定のボタンを押下するようにする。自動通訳サーバ1000中の音声生成手段12は、電話端末1上の次候補提示の機能を割り当てた所定のボタンが押下されると、メモリ15上のJCNT208をインクリメントし、RECOGS (JCNT)を読み出し、該文字列を合成音声に変換し、該音声の波形データをデジタル／アナログ変換によりアナログデータに変換し、音声出力手段13を通じて音声として電話端末1に送る。

【0035】

上述した動作と同様に、ユーザは、前記送られた音声を電話端末1上のスピーカ100から聞き、該音声が自分の発声入力した内容に合致していれば、電話端末1上の確定の機能を割り当てた所定のボタンを押下し確定し、該音声が自分の発声入力した内容に合致していなければ、ユーザは電話端末1上の次候補提示の機能を割り当てた所定のボタンを押下し、自分の発声入力した内容に合致する音声聞こえるまで、以上の動作を繰り返す。

【0036】

確定ボタンを押下すると、メモリ15上のRECOGS (ICNT)に格納されている文字列を同じくメモリ15上のJAPANESE 203に格納する。この確定ボタンを押下する代わりに、ある所定の単語もしくはフレーズもしくは文をユーザが発声入力する方法を用いても良い。すなわち、ユーザは、前記送られた音声を電話端末1上のスピーカ100から聞き、該音声が自分の発声入力した内容に合致していれば、合致している旨の所定の単語もしくはフレーズもしくは文を確定音声として電話端末1上のマイク104から発声入力し、該音声が自分の発声入力した内容に合致していなければ、合致している場合とは異なる所定の単語もしくはフレーズもしくは文を非確定音声として電話端末1上のマイク104から発声入力するようにし、自動通訳サーバ1000中の音声認識手段6が、上述したような文入力と同様の方法で音声を認識し、全ての尤度があらかじめ設定した閾値を下回る場合もしくはICNT204

の値がNを超えた場合、コマンド辞書21との照合を開始する。コマンド辞書21のデータ構造を図16に示す。入力音声の特徴ベクトルの時系列を、コマンド辞書21の各項目の特徴ベクトルの時系列と照合していき、尤度の最も高い候補のコマンド番号をもってコマンドとする。ここでは例として、ユーザが「かくてい」と発声入力した場合を想定すると、まず、文辞書20を用いた認識で、該音声の特徴ベクトルと各項目の特徴ベクトルの照合で尤度があらかじめ設定した閾値を下回ることになり、コマンド辞書21の各項目の特徴ベクトルの照合が行なわれ、701の項目が認識候補となる。該項目のコマンド番号が1であれば、それは「確定」を表す発声入力ということになる。

【 0 0 3 7 】

確定音声であれば、メモリ15上のRECOGS(ICNT)に格納されている文字列を同じくメモリ15上のJAPANESE 203に格納する。また、非確定音声であれば、メモリ15上のJCNT208をインクリメントし、RECOGS(JCNT)を読み出し、該文字列を合成音声に変換し、該音声の波形データをデジタル／アナログ変換によりアナログデータに変換し、音声出力手段13を通じて音声として電話端末1に送るという上述の動作を確定音声となるまで繰り返すようにする。

【 0 0 3 8 】

次に、自動通訳サーバ1000中の言語翻訳手段7が起動される。言語翻訳手段7は、言語変換テーブル10を用いてメモリ上のJAPANESE 203に格納されている文字列を別の言語に翻訳変換する。以下、言語翻訳手段7の動作について説明する。言語変換テーブル10のデータ構造を図9に示す。

【 0 0 3 9 】

まず、言語翻訳手段7は、メモリ15上のPCAND213、SCAND214に格納されている値を言語変換テーブル10の文例番号311および文番号312の各項目と順次照合していき、一致した項目の訳語312の中のLANG209の値のカラムの内容をメモリ15のRESULT206に格納する。

【 0 0 4 0 】

先の例では、PCAND213、SCAND214に格納されている値は各々”4”、”1”であるので、303の項目”みずがほしいです”と一致する。さらに、LANG209の値は”2”で

あるので、対応する訳語312の”I'd like to have a glass of water”をメモリ15のRESULT 206に格納する。訳語への変換については、上述した例のような言語変換テーブルを用いる方法のみに限らず、例えば特開平03-328585に記載されている用例データベースを用いた翻訳方法や、特開平03-51022に記載されている構文解析による翻訳方法などを使用してもよい。

【0041】

次に、言語翻訳手段7は、図10に示すように、メモリ15のJAPANESE 203およびRESULT 206に各々格納されている内容を電話端末1上のディスプレイ101に表示する。同図の表示方法は一例であって、表示の仕方が本発明の範囲を限定するものではない。

【0042】

続いて、自動通訳サーバ1000中の音声生成手段12が起動される。音声生成手段12はメモリ15上のENGLISH 206に格納されている文字列を読み込み、該文字列を合成音声に変換し、メモリ15上のSYNWAVE 207に波形データを格納する。文字列を合成音声に変換するためには、例えば、J.Allen, M.S.Hunnicutt and D.Klatt他”From Text to Speech”(Cambridge University Press, 1987) P.16からP.150に記載されているホルマント合成方式や、矢頭「テキスト音声合成技術の最新状況」(Interface, Dec., 1996) P.161からP.165に記載されている波形重畳方式が使える。もちろん、他のテキスト音声合成方式を使うこともできる。また、各英訳に対応する音声をあらかじめ録音して、そのままあるいは圧縮してメモリ等の記憶装置に格納しておくようにしてもよい。

【0043】

続いて、音声生成手段12は、メモリ15上のSYNWAVE 207に格納された訳語音声の波形データをアナログデータもしくはパケットデータに変換し、音声出力手段13を通じて音声として電話端末1に送り、該電話端末1の有するメモリに前記送られた訳語音声を格納する。ただし、音声出力手段13から出力された訳語音声を前記電話端末1の有するメモリに格納しないで、自動通訳サーバ1000のメモリ15上に格納するようにしても良い。

【 0 0 4 4 】

この時点で、課金テーブル33のユーザ I D カラム411の前記認証サーバ31から送られた I D セルに対応する料金カラム412の内容に通訳 1 回分の所定の料金が加えられる。例えば通訳一回あたり50円という料金を設定してあれば、先の例では、ユーザ I D が”1236”であるので、課金テーブル33のユーザ I D カラム411のセルのうち”1236”であるセル403に対応する料金カラム412のセルが”50”となる。この料金は、1 通訳サービスあたりの金額で設定しても良いし、ある定額の中で何度も通訳サービスを利用できるように設定しても良い。サービス終了後、加算された料金は、各 I D のユーザに対し請求され支払いが行なわれる。料金のうちの手数料などとして所定の金額がモバイルインターネットアクセスサービスゲートウェイサーバ3の所有者に支払われ、残りの金額が自動通訳サーバ1000の所有者に支払われるようにする。ただし、これらの料金の分配方法や分配先は一例であり、状況に応じて設定することができる。

【 0 0 4 5 】

ここで、ユーザが電話端末1上の訳語出力の機能を割り当てた所定のボタンを押下すると、該電話端末1上のメモリに格納されている訳語音声を読み出し、該電話端末1上のスピーカ100より該訳語音声出力されるようにする。さらに、ユーザが電話端末1上の訳語出力の機能を割り当てた所定のボタンを押下すると、そのたびに該電話端末1上のメモリに格納されている訳語音声を読み出し、電話端末1上のスピーカ100より該訳語音声出力されるようにする。

【 0 0 4 6 】

該訳語音声出力するための方法には、電話端末1上の訳語出力の機能を割り当てた所定のボタンを押下する方法に限らず、所定の単語もしくはフレーズもしくは文をユーザが発声入力する方法を用いても良い。すなわち、ユーザは、前記訳語音声の出力を指示する旨の所定の単語もしくはフレーズもしくは文を出力指示音声として電話端末1上のマイク104から発声入力し、自動通訳サーバ1000中の音声認識手段6が該音声を認識し、出力指示音声であれば、該電話端末1上のメモリに格納されている訳語音声を読み出し、電話端末1上のスピーカ100より該訳語音声出力されるようにする。ただし、音声出力手段13から出力された訳語音声

前記電話端末1の有するメモリに格納されていない場合は、自動通訳サーバ1000のメモリ15上に格納されている訳語音声の前記電話端末1に送るようにする。

【0047】

なお、ここで述べた前記電話端末1上のメモリに格納されている訳語音声を読み出し電話端末1上のスピーカ100より該訳語音声が出力させる場合には、モバイルインターネットアクセスサービスゲートウェイサーバ3に何ら情報を送ることはなく、したがって課金サーバ32によって前記ユーザに課金がなされることはない。以上の動作の流れを図18、図19にフローチャートで示す。

【0048】

続いて、本発明の第2の実施例について図を用いて説明する。図17は本発明の第2の実施例を示す自動通訳サービスシステム構成図である。第1の実施例と同様に、本実施例におけるサーバは、パソコン、ワークステーション等、中央演算装置、メモリを備え、同図のような構成をなすことができるものであれば何でもよく、プラットフォームの種類が本発明の適用範囲を限定するものではない。また、本実施例は英語を日本語に通訳する自動通訳サービスであるが、これは一例であり、日英、中日、日中等、言語の種類は限定されない。同図において、1はモバイルインターネットアクセスサービス利用可能な電話端末、2はモバイルインターネットアクセスサービスパケット網、3はモバイルインターネットアクセスサービスゲートウェイサーバ、4は電話回線ボード、5は音声入力手段、6は音声認識手段、7は言語翻訳手段、8は単語辞書、9は文法テーブル、10は言語変換テーブル、11は音声生成手段、12は素片セット、13は音声出力手段、14は中央演算装置、15はメモリ、16は言語種表示手段、17は場面表示手段、18は文例表示手段、19は認識候補表示手段、20は文辞書、21はコマンド辞書、22は言語種テーブル、23は場面テーブル、24は文例テーブル、31は認証サーバ、32は課金サーバ、33は課金テーブル、34は通話用通信網、1000は自動通訳サーバ、1001はモバイルインターネットアクセスサービスパケット網に接続される回線、1002は通話用通信網に接続される回線である。

【0049】

メモリ15のデータ構造及びモバイルインターネットアクセスサービス利用可能な電

話端末1の概観の例は第1の実施例と同様である。ただし、ここでは、第1の実施例とは異なり、モバイルインターネットアクセスサービス利用可能な電話端末として、通話音声およびデータを同一のプロトコルで扱えない電話端末を仮定し説明する。

【0050】

図17、図3において、まず、モバイルインターネットアクセスサービス利用可能な電話端末1の電源104をいれ、モバイルインターネット接続用ボタン102を押下すると、モバイルインターネットアクセスサービスパケット網2を介して、センタのモバイルインターネットアクセスサービスゲートウェイサーバ3に接続され、認証サーバ31でモバイルインターネットアクセスサービス利用登録ユーザであることが確認される。以下、場面確定までの動作は、第1の実施例と同様である。

【0051】

場面が確定すると、次に、自動通訳サーバ1000の文例表示手段18が起動され、文例テーブル24を用いて回線1001を通じて電話端末1上のディスプレイ101に、図7のように通訳可能な文例を表示すると同時に、自動通訳サーバ1000の音声入力手段5が起動される。続いて、音声入力手段5は、システムを音声入力可能な状態にする。その後、ユーザは、該文例を見ながら、例えば先の例であれば、日本語でレストランの場面での通訳させたい文を電話端末1の送話口のマイク104に向かって発声入力する。ここでは、例として、レストランの場面で「水が欲しいです。」という文を英語に通訳したい場合を考えることにする。文例テーブル24は図14のようなデータ構造をもっており、文例表示手段18は、文例テーブル24の場面番号511の各項目のうち、メモリ15上のSECENE210に格納されている値の項目の文例513を文例番号512の”1”から順次、前記電話端末1に送り該電話端末1上のディスプレイ101に表示する。先の例では、メモリ15上のSCENE210に”3”が格納されているので、図14の文例テーブル24の場面番号511の値が”3”である項目501,502,503,504の文例513の内容「こんにちは」「ありがとう」「[]はどこですか」「[]が欲しいです」を前記電話端末1に送り図7のように該電話端末1上のディスプレイ101にM個ずつ順次表示する。Mはディスプレイ101の大きさに応じて設定する定数であるが、ここではM=4とする。

【0052】

ユーザは、前記文例の中に「[]が欲しいです」という文例があるので、この文例を基に「水が欲しいです」と発声入力する。この発声入力に際して、電話端末1上の発声入力の合図の意味をもたせた所定のボタンを押下して自動通訳サーバ1000の音声入力手段5が音声入力可能な状態にしてから発声するようにしても良いし、特にボタン等のトリガを用いなくて自動通訳サーバ1000の音声入力手段5が起動された後いつでも音声入力可能な状態にしておいても良い。なお、表示する文例は、上述の例のようにスロット[]を含有した文もしくは文法でも良いし、もしくは文そのものでも良い。

【0053】

第2の実施例では、通話音声およびデータを同一のプロトコルで扱えない電話端末を仮定しているため、発声入力した音声を自動通訳サーバに送るためには、通話用通信網20を介する必要がある。したがって、ユーザは、前記電話端末1から、モバイルインターネットアクセスサービスゲートウェイサーバ3経由で回線1001を通じて自動通訳サーバ1000に接続したのとは別の電話番号で通話用通信網34経由で回線1002を通じて自動通訳サーバに接続する。ここで、ユーザが改めて接続する方法の他に、自動通訳サーバ1000の音声入力手段5が前記ユーザの電話端末1に接続するようにしても良い。いずれにしても、ユーザの発声した音声は、通話用通信網34経由で回線1002を通じて、自動通訳サーバ1000に送られる。

【0054】

以下、音声認識手段6が起動され前記発声入力された音声認識され、結果がメモリに格納される動作までは、第1の実施例と同様である。次に、認識候補表示手段19が起動され、RECOGS(1)、RECOGS(2)、…、RECOGS(M)205の内容を、図8のように、回線1001を通じて電話端末1に送り該電話端末1上のディスプレイ101に順次表示する。このとき、メモリ15上のICNT204に"1"を格納し、電話端末1上のディスプレイ101に表示されているRECOGS(ICNT)の内容が白黒反転されるようにする。Mはディスプレイ101の大きさに応じて設定する定数であるが、ここではM=4とする。また、メモリ15上のINDEX215に"0"を格納する。

【0055】

ユーザは、第1候補の表示が自分の発声した内容と一致しているもしくは近ければ、そのまま確定の機能を割り当てた所定のボタンを押下し確定する。第1候補の表示が正しくない場合、ユーザは、上下カーソル移動の機能を割り当てた所定のボタンを押下することによって、電話端末1上のディスプレイ101上の正しい文字列が表示されている部分にまで白黒反転表示領域を下方に移動させる。すなわち、下移動ボタンを押下するたびに、メモリ15上のICNT 204の値をインクリメントし、回線1001を通じて電話端末1上のディスプレイ101上の同じくメモリ15上のRECOG(ICNT)の内容が表示されている部分のみを白黒反転表示するようにする。

ここで、ICNT 204の値がMを超えた場合は、メモリ15上のINDEX215の値に”M”を加算し、次のM個の候補RECOGS(INDEX+1)、RECOGS(INDEX+2)、…、RECOGS(INDEX+M)をメモリ15上の中から読み出し、回線1001を通じて電話端末1に送り該電話端末1上のディスプレイ101に順次表示する。このとき、メモリ15上のICNT 204に”1”を格納し、電話端末1上のディスプレイ101に表示されているRECOGS(INDEX+1)、RECOGS(INDEX+2)、…、RECOGS(INDEX+M)のうちのICNT番目の表示が白黒反転されるようにする。以下、同様にして、次のM個の候補を、電話端末1に送り該電話端末1上のディスプレイ101に順次表示していく。なお、図15に示した音声認識に使用する文辞書20の構造は一例であり、例えば、文法と単語辞書を組み合わせる方法もある。その場合も、メモリ15上のSCENE210に格納された場面番号に対応させて、場面別の文法や辞書を用いるようにする。

【0056】

自分の発声した内容の確認動作は、第1の実施例と同様であるが、自動通訳サーバ1000中の音声生成手段12は、電話端末1上の次候補提示の機能を割り当てた所定のボタンが押下されると、メモリ15上のJCNT208をインクリメントし、RECOGS(JCNT)を読み出し、該文字列を合成音声に変換し、該音声の波形データをデジタル／アナログ変換によりアナログデータに変換し、音声出力手段13を通じて音声として回線1001を介して電話端末1に送る。

【0057】

確定ボタンを押下すると、メモリ15上のRECOGS(ICNT)に格納されている文字列

を同じくメモリ15上のJAPANESE 203に格納する。確定ボタンの信号を自動通訳サーバ1000に送るルートは、回線1001経由でもよいし、回線1002経由でもよい。この確定ボタンを押下する代わりに、ある所定の単語もしくはフレーズもしくは文をユーザが発声入力する方法を用いても良い。すなわち、ユーザは、前記送られた音声を電話端末1上のスピーカ100から聞き、該音声が自分の発声入力した内容に合致していれば、合致している旨の所定の単語もしくはフレーズもしくは文を確定音声として電話端末1上のマイク104から発声入力し、該音声が自分の発声入力した内容に合致していなければ、合致している場合とは異なる所定の単語もしくはフレーズもしくは文を非確定音声として電話端末1上のマイク104から発声入力するようにし、回線1002を通じて自動通訳サーバ1000に送られ、該自動通訳サーバ中の音声認識手段6が、上述したような文入力と同様の方法で音声を認識し、全ての尤度があらかじめ設定した閾値を下回る場合もしくはICNT204の値がNを超えた場合、コマンド辞書21との照合を開始する。次に、自動通訳サーバ1000中の言語翻訳手段7が起動されるが、該言語翻訳手段7の翻訳動作は、第1の実施例と同様である。

【0058】

次に、言語翻訳手段7は、図10に示すように、メモリ15のJAPANESE 203およびRESULT 206に各々格納されている内容を回線1001を通じて電話端末1上のディスプレイ101に表示する。同図の表示方法は一例であって、表示の仕方が本発明の範囲を限定するものではない。

【0059】

続いて、自動通訳サーバ1000中の音声生成手段12が起動されるが、該音声生成手段12が音声を生成する動作は第1の実施例と同様である。

【0060】

続いて、音声生成手段12は、メモリ15上のSYNWAVE 207に格納された訳語音声の波形データをアナログデータもしくはパケット音声に変換し、音声出力手段13を通じて音声として回線1002を通じて電話端末1に送り、該電話端末1の有するメモリに前記送られた訳語音声を格納する。以下、課金の方法及び動作についても、第1の実施例と同様である。

【 0 0 6 1 】

【発明の効果】

本発明によれば、携帯電話端末を用いた精度の高い自動通訳サービスが実現できるので、異国人同士が互いに相手の言語を意識することなく容易にコミュニケーションを行なうことが可能となる。

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施例を示す日英自動通訳サービスシステム構成図。

【図 2】

メモリ 15 のデータ構造の例。

【図 3】

電話端末 1 の例。

【図 4】

電話端末 1 のディスプレイ 101 へのサービスメニュー表示例。

【図 5】

電話端末 1 のディスプレイ 101 への通訳可能な言語種表示例。

【図 6】

電話端末 1 のディスプレイ 101 への通訳可能な場面表示例。

【図 7】

電話端末 1 のディスプレイ 101 への通訳可能な文例表示例。

【図 8】

電話端末 1 のディスプレイ 101 への認識結果表示例。

【図 9】

言語変換テーブル 10 のデータ構造の例。

【図 1 0】

電話端末 1 のディスプレイ 101 への通訳結果表示例。

【図 1 1】

課金テーブル 33 のデータ構造の例。

【図 1 2】

言語種テーブル22のデータ構造の例。

【図 1 3】

場面テーブル23のデータ構造の例。

【図 1 4】

文例テーブル24のデータ構造の例。

【図 1 5】

文辞書20のデータ構造の例。

【図 1 6】

コマンド辞書21のデータ構造の例。

【図 1 7】

本発明の第 2 の実施例を示す日英自動通訳サービスシステム構成図。

【図 1 8】

第 1 の実施例の動作を示すフローチャート（その 1）。

【図 1 9】

第 1 の実施例の動作を示すフローチャート（その 2）。

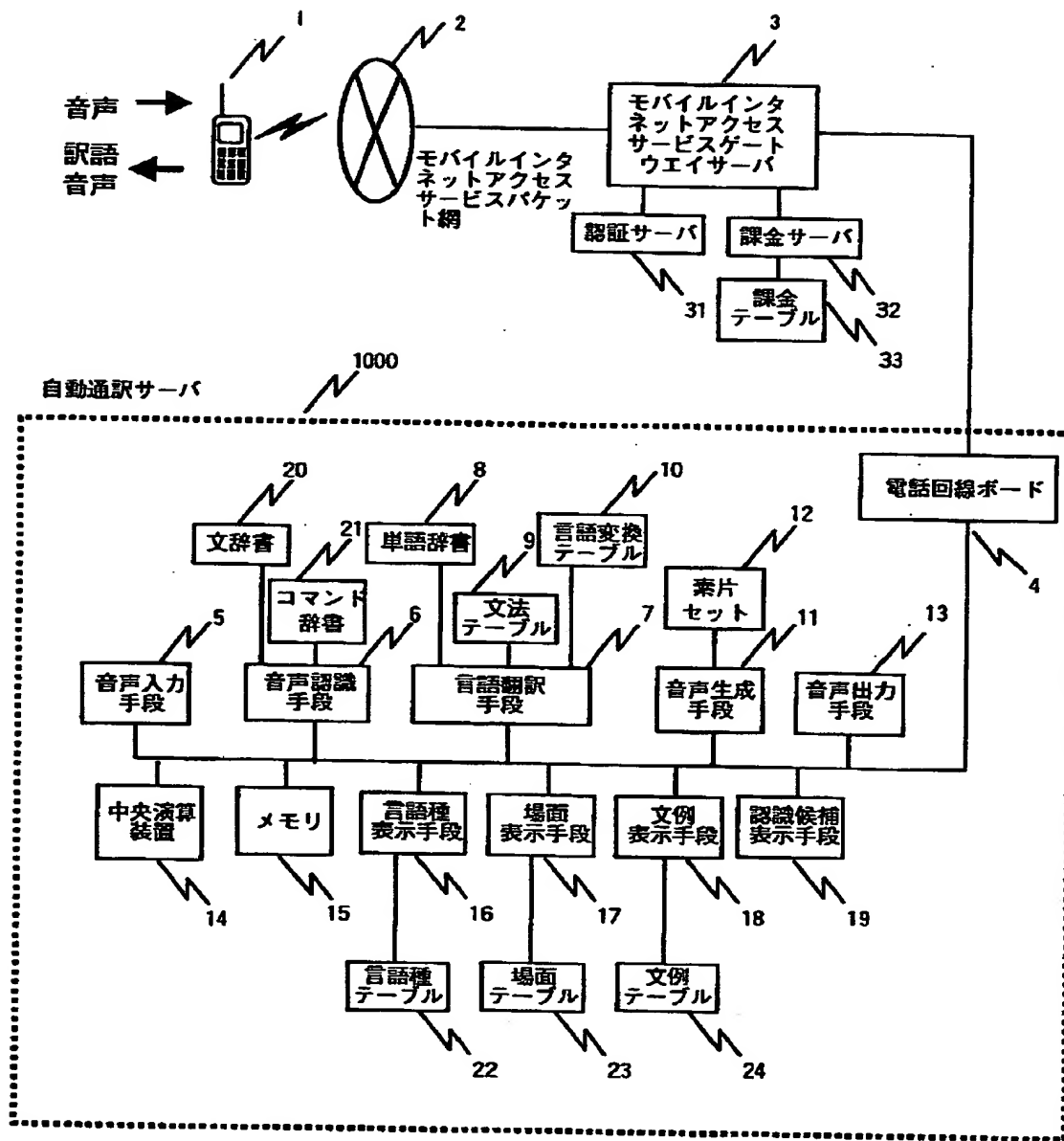
【符号の説明】

1はモバイルインタネットアクセスサービス利用可能な電話端末、2はモバイルインタネットアクセスサービスパケット網、3はモバイルインタネットアクセスサービスゲートウェイサーバ、4は電話回線ボード、5は音声入力手段、6は音声認識手段、7は言語翻訳手段、8は単語辞書、9は文法テーブル、10は言語変換テーブル、11は音声生成手段、12は素片セット、13は音声出力手段、14は中央演算装置、15はメモリ、16は言語種表示手段、17は場面表示手段、18は文例表示手段、19は認識候補表示手段、31は認証サーバ、32は課金サーバ、33は課金テーブルである。

【書類名】 図面

【図 1】

図 1



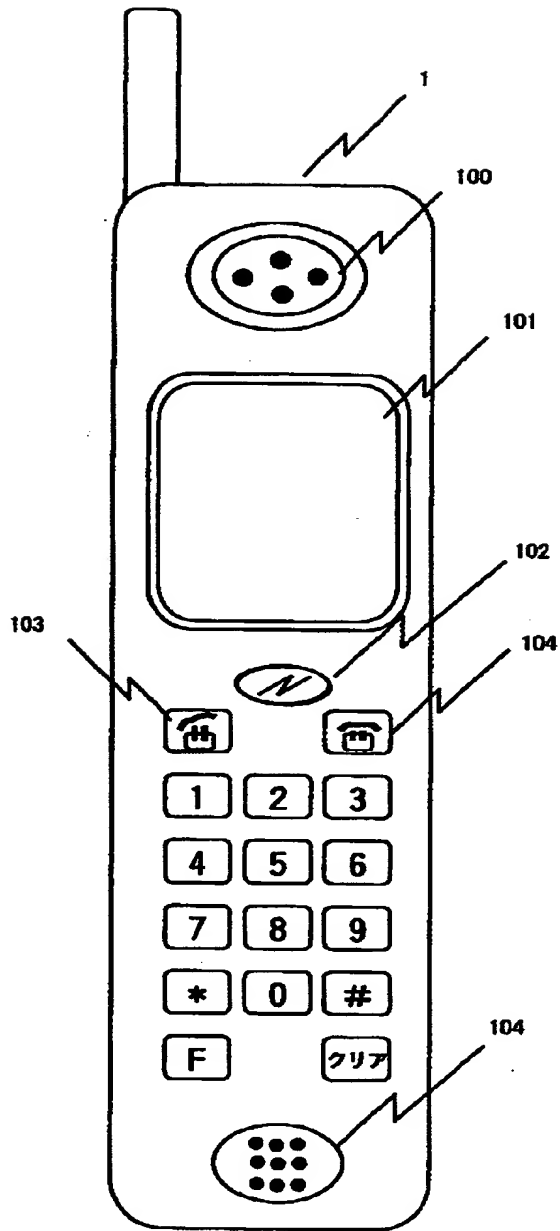
【図 2】

図 2

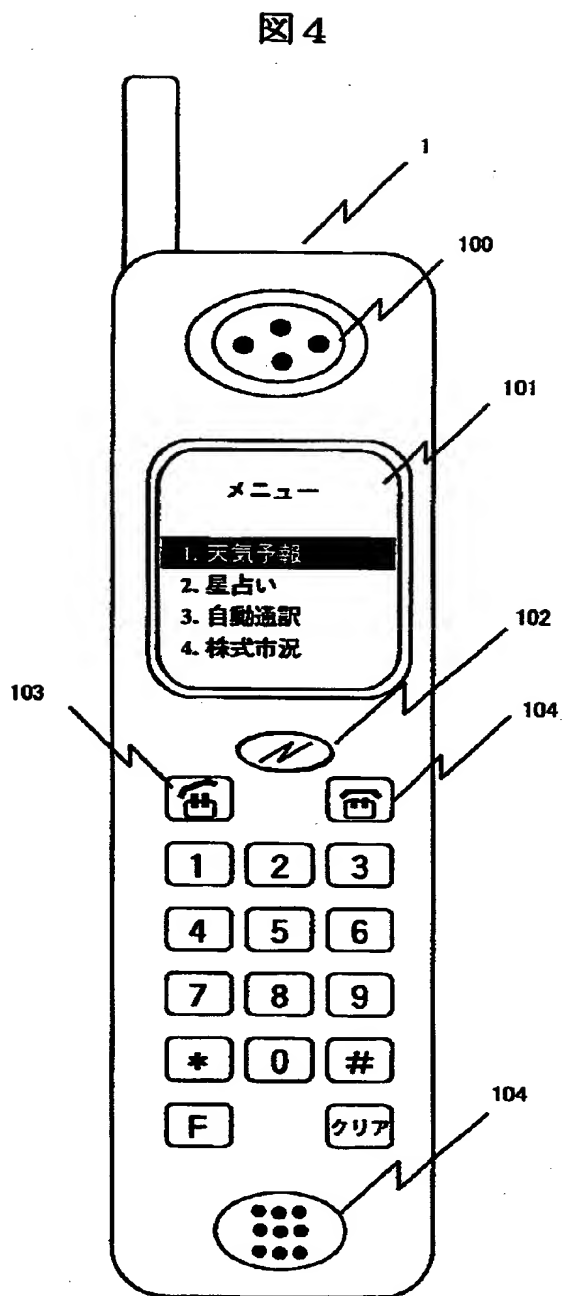
| | | |
|-----|--------------|--|
| 201 | WAVE | |
| 202 | CSTRING | |
| 203 | JAPANESE | |
| 204 | ICNT | |
| 205 | RECOGS(1) | |
| | RECOGPNUM(1) | |
| | RECOGSNUM(1) | |
| | : | |
| | RECOGS(N) | |
| | RECOGPNUM(N) | |
| | RECOGSNUM(N) | |
| 206 | RESULT | |
| 207 | SYNRAVE | |
| 208 | JCNT | |
| 209 | LANG | |
| 210 | SCENE | |
| 211 | BSSENTENCE | |
| 212 | CSSENTENCE | |
| 213 | PCAND | |
| 214 | SCAND | |
| 215 | INDEX | |
| | : | |

【図 3】

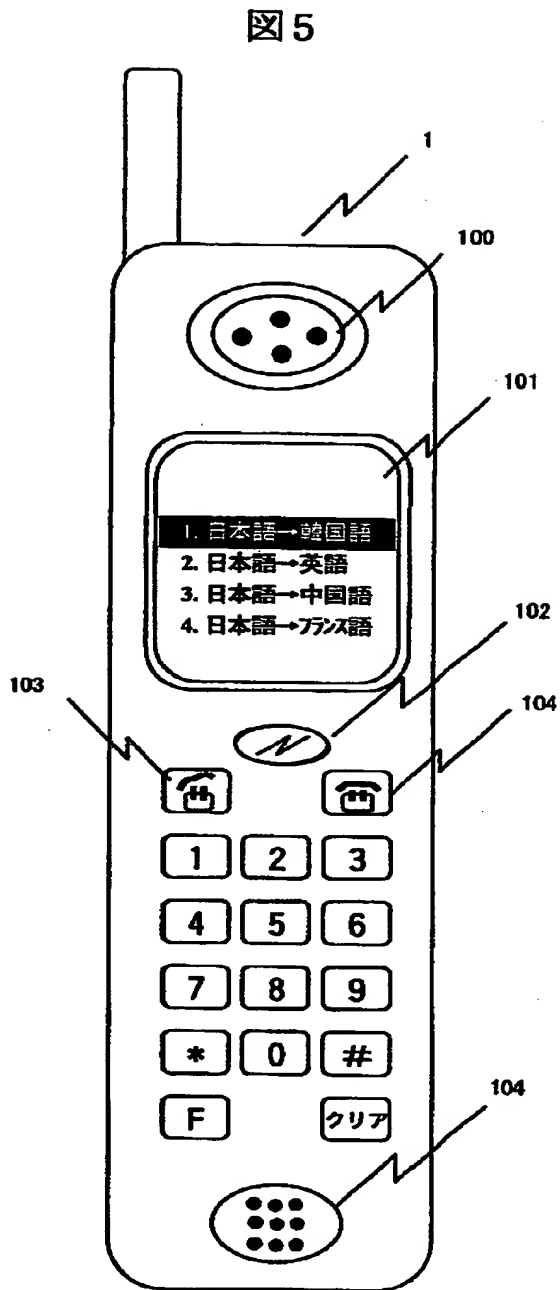
図 3



【図4】

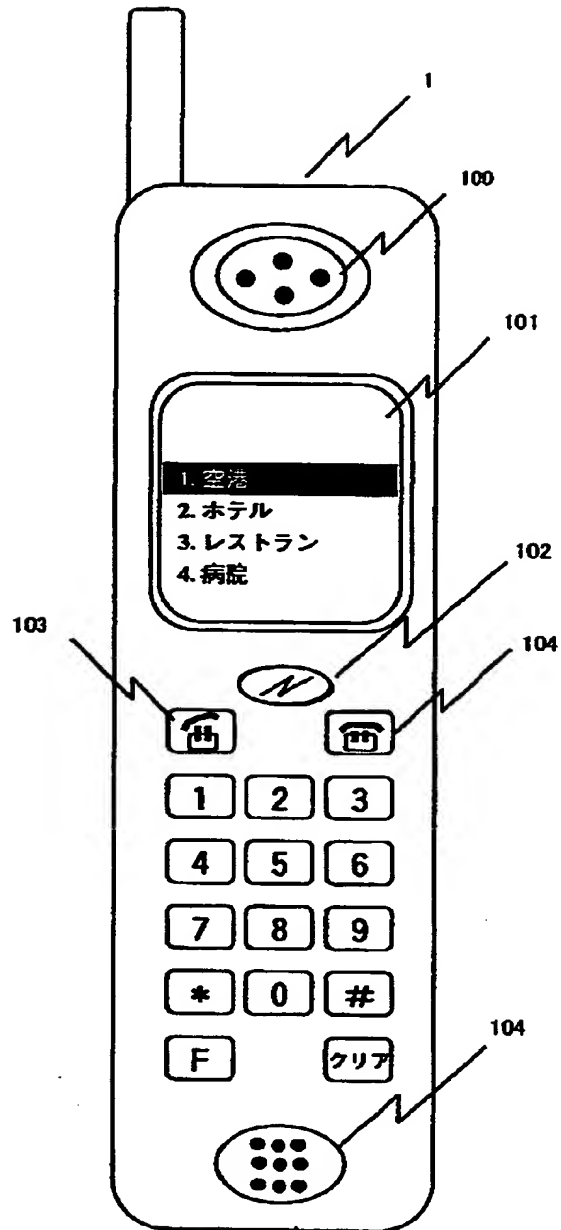


【図 5】



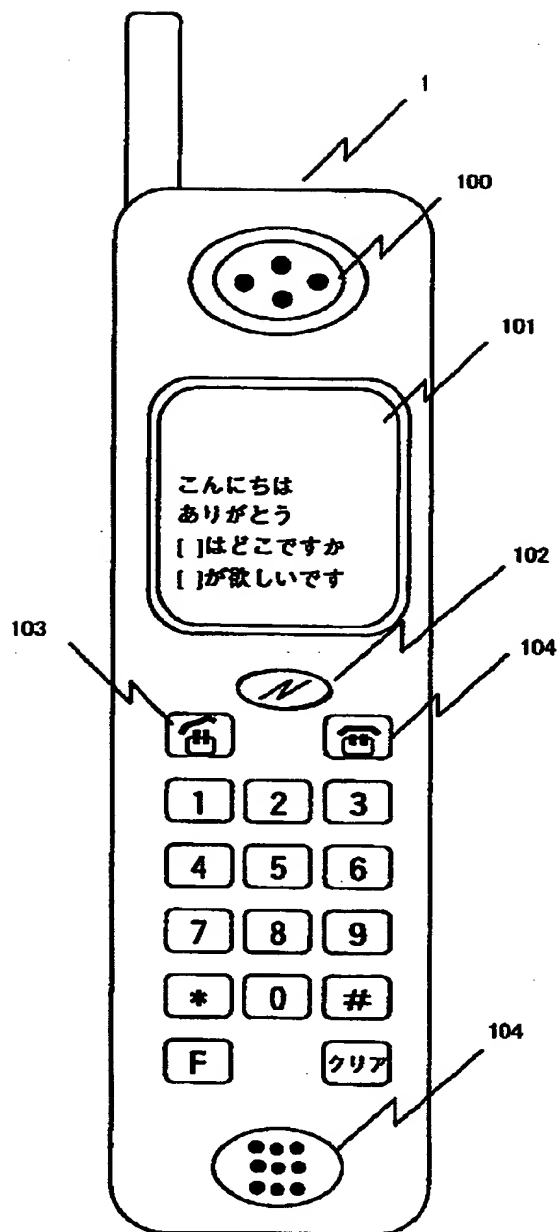
【図 6】

図 6

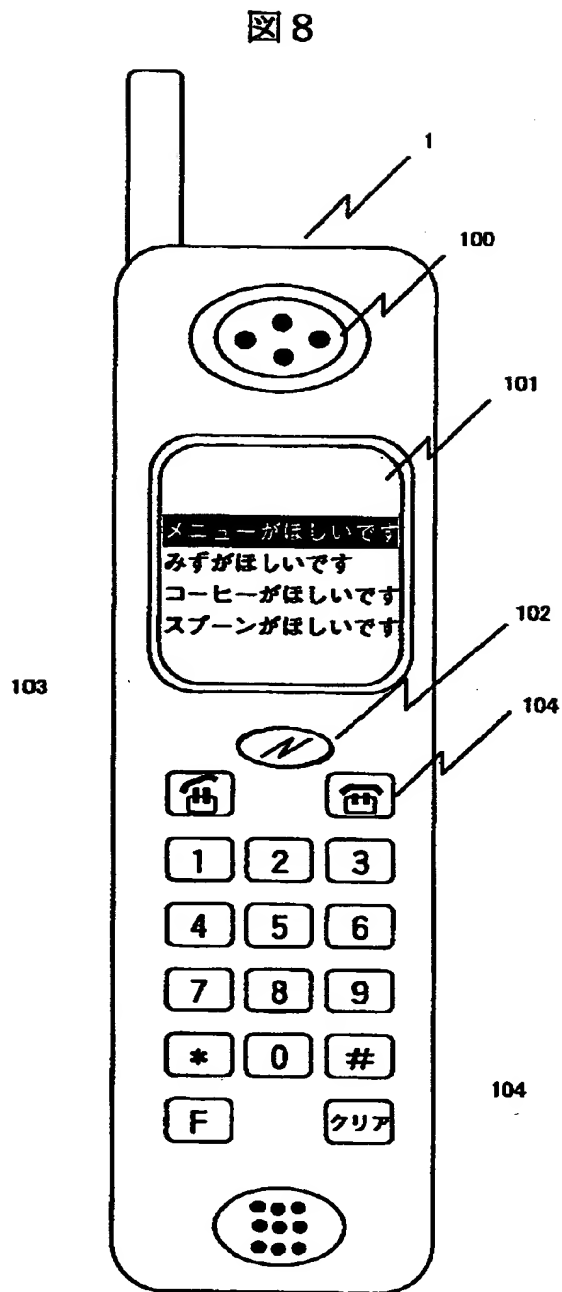


【図 7】

図 7



【図 8】



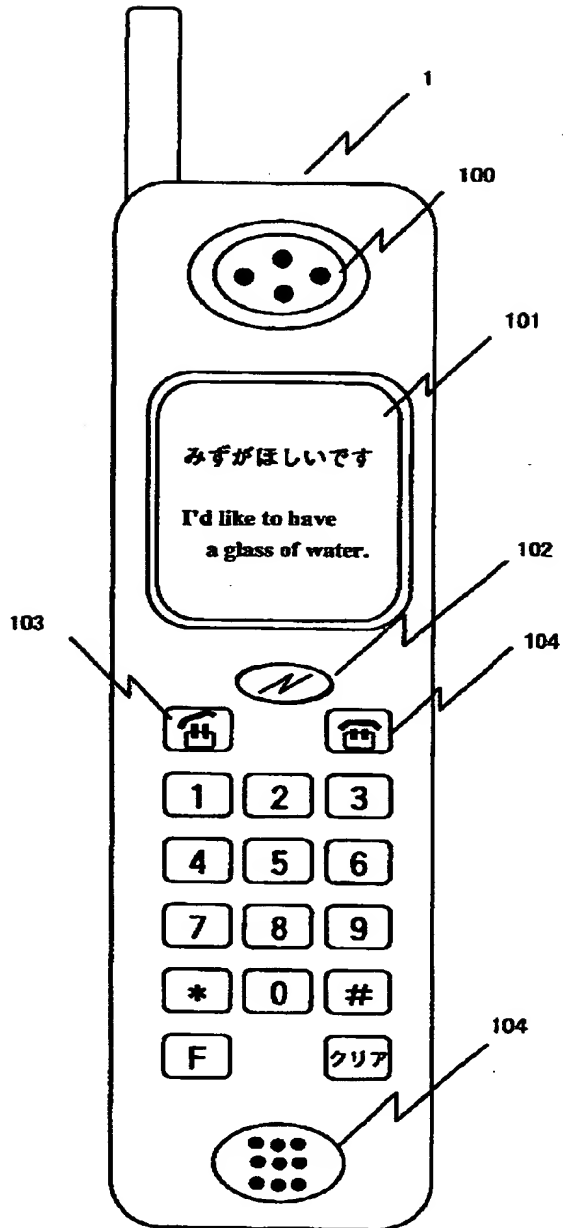
【図 9】

図 9

| 文例番号 | 文番号 | 訳語 | | |
|------|-----|-----|------------------------------------|-----|
| | | 1 | 2 | ... |
| : | : | : | : | : |
| 4 | 1 | ... | I'd like to have a glass of water. | ... |
| 4 | 2 | ... | Can I see the menu? | ... |
| 4 | 3 | ... | I'd like to have a cup of coffee. | ... |
| 4 | 4 | ... | I'd like to have a spoon. | ... |
| : | : | : | : | : |

【図10】

図10



【図11】

図11

| | | | | |
|-----|-----|-------|-----|-----|
| | 411 | ユーザID | 412 | 料金 |
| | | : | | : |
| 401 | | 1234 | | 100 |
| 402 | | 1235 | | 0 |
| 403 | | 1236 | | 300 |
| 404 | | 1237 | | 850 |
| | | : | | : |

【図12】

図12

| | | | | |
|-----|-----|-------|-----|-----------|
| | 811 | 言語種番号 | 812 | 言語種 |
| 801 | | 1 | | 日本語→韓国語 |
| 802 | | 2 | | 日本語→英語 |
| 803 | | 3 | | 日本語→中国語 |
| 804 | | 4 | | 日本語→フランス語 |
| | | : | | |

【図 13】

図 13

| | | | |
|-----|------|-------|-----|
| | 911 | | 912 |
| 901 | 場面番号 | 場面 | |
| 902 | 1 | 空港 | |
| 903 | 2 | ホテル | |
| 904 | 3 | レストラン | |
| | 4 | 病院 | |

【図 14】

図 14

| | | | |
|-----|------|------|-----------|
| | 511 | 512 | 513 |
| | 場面番号 | 文例番号 | 文例 |
| | 1 | 1 | []はどこですか |
| | : | : | |
| 501 | 3 | 1 | こんにちは |
| 502 | 3 | 2 | ありがとう |
| 503 | 3 | 3 | []はどこですか |
| 504 | 3 | 4 | []が欲しいです |
| 505 | 3 | 5 | []はありますか |
| 506 | 3 | 6 | []てください |
| | : | : | |

【図 15】

図 15

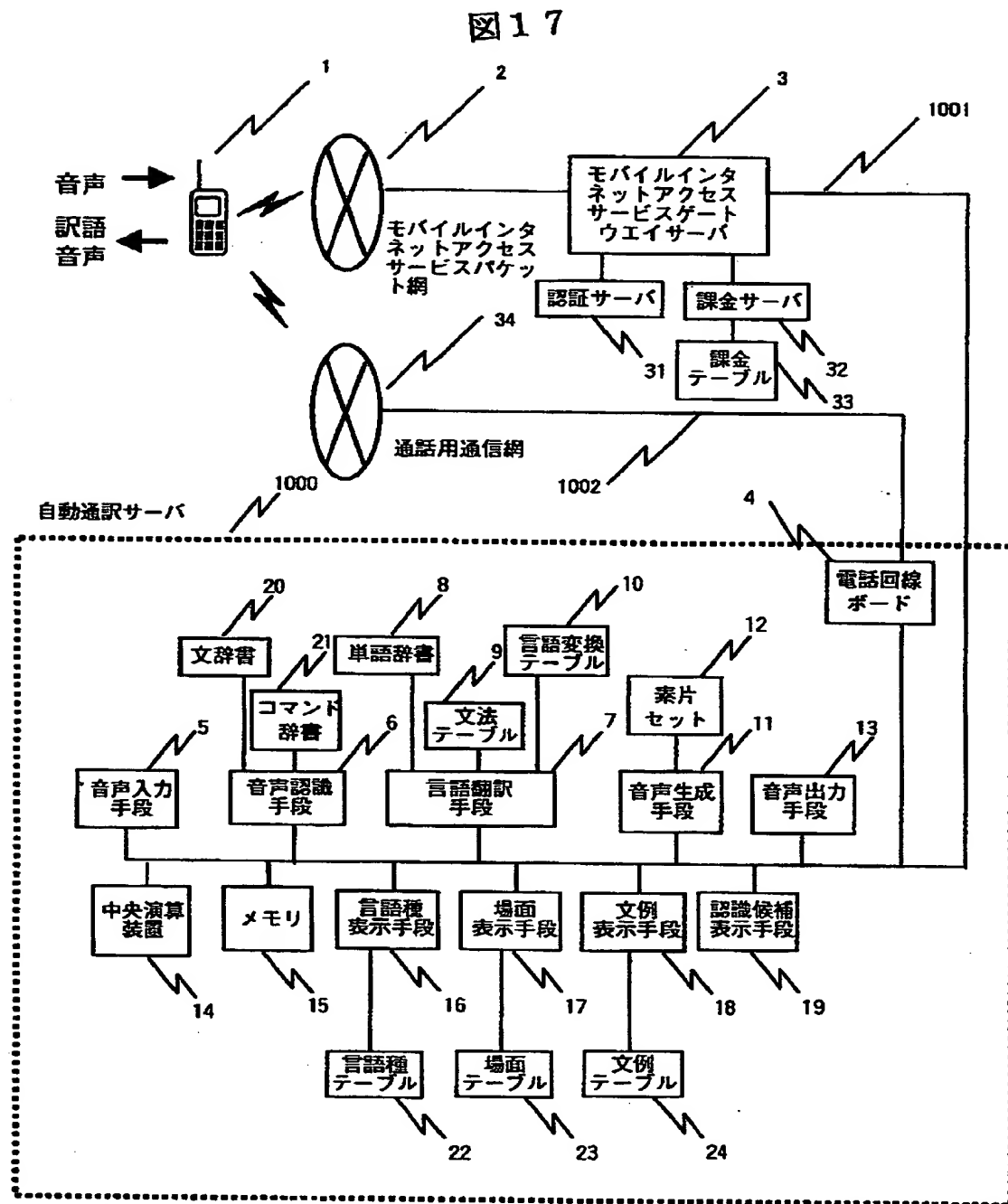
| 601 | 文例番号 | 文番号 | 文 | 特徴ベクトル時系列 |
|-----|------|-----|------------|---------------|
| 602 | 1 | 1 | こんにちは | A11, A12, ... |
| 603 | 2 | 1 | ありがとう | A21, A22, ... |
| 604 | : | : | : | : |
| 605 | 4 | 1 | みずがほしいです | A31, A32, ... |
| 606 | 4 | 2 | めにゅーがほしいです | A41, A42, ... |
| | 4 | 3 | こーひーがほしいです | A51, A52, ... |
| | 4 | 4 | すぶーんがほしいです | A61, A62, ... |
| | : | : | : | : |

【図 16】

図 16

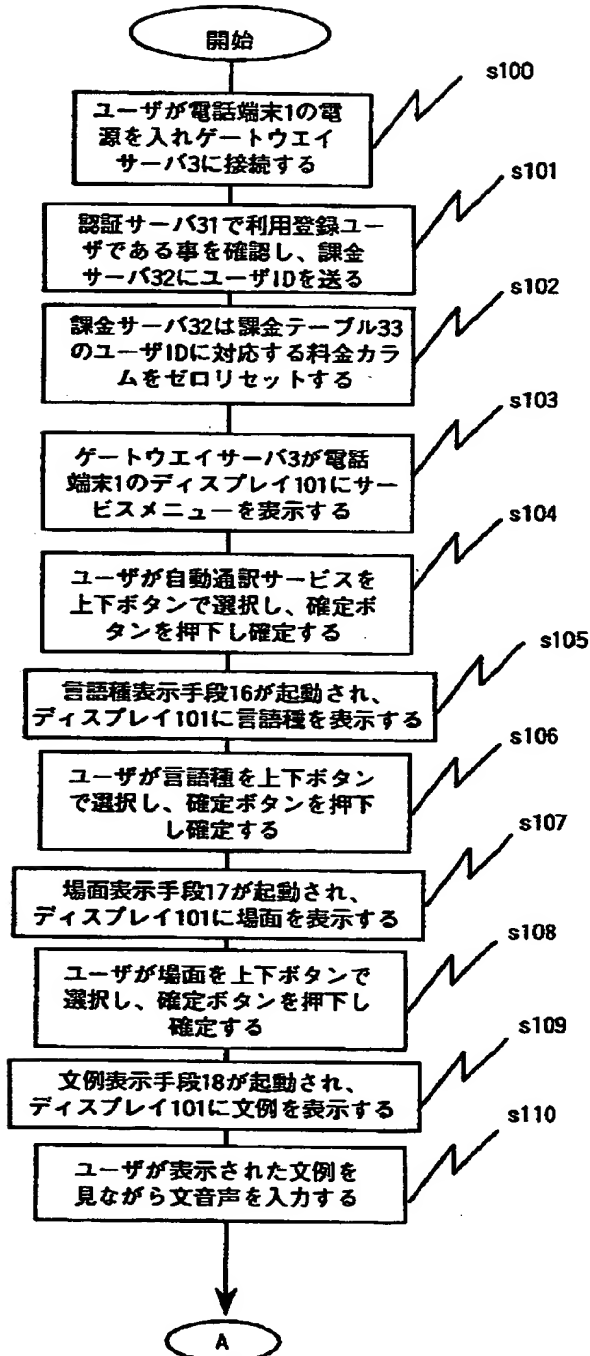
| 701 | コマンド番号 | コマンド | 特徴ベクトル時系列 |
|-----|--------|--------|---------------|
| 702 | 1 | かくてい | B11, B12, ... |
| 703 | 2 | つぎ | B21, B22, ... |
| | 3 | サービス終了 | : |
| | : | | B31, B32, ... |

【图 1-7】



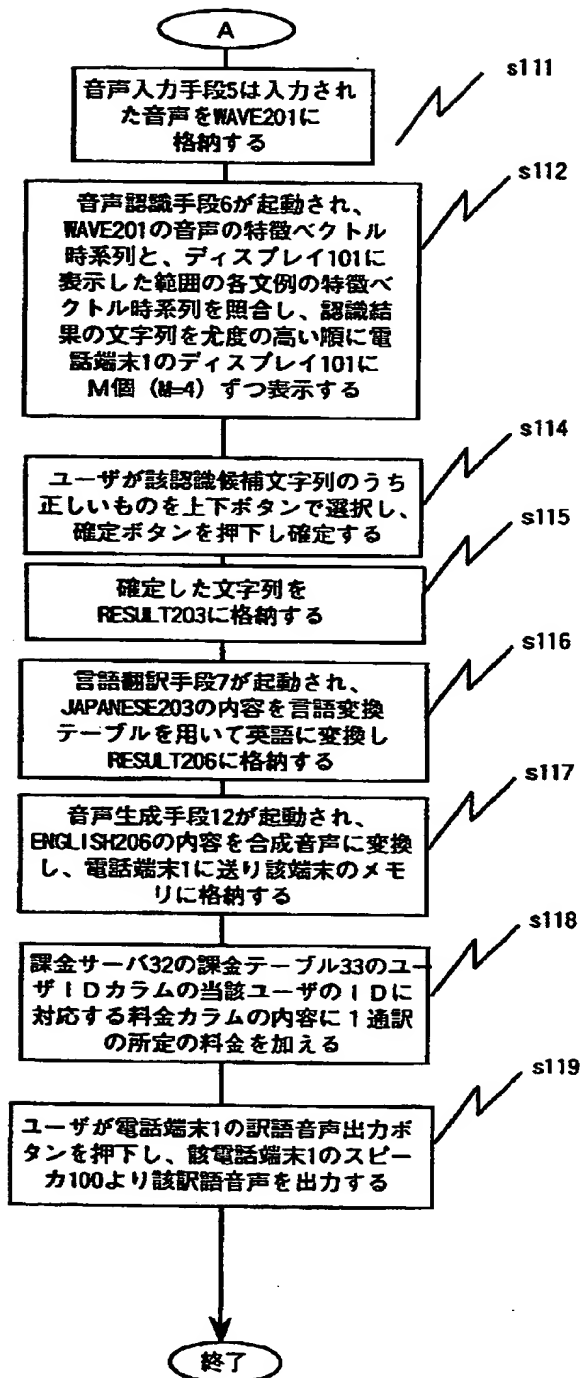
【図 18】

図 18



【図19】

図 19



【書類名】 要約書

【要約】

【課題】 本発明の目的は、通話用の電話機、とりわけ、普及率の高いモバイルインターネットアクセスサービス利用可能な電話を用いて通訳精度の高い音声通訳が誰でも簡単に利用できるサービスを提供することにある。

【解決手段】 上記課題を解決するために、本発明では、ユーザが電話を用いてサーバ上で稼動する自動通訳サーバに音声を入力、送信すると、翻訳された音声を、再びユーザの電話に返信するようにする。その際に、最初にユーザがモバイルインターネットアクセスサービス利用可能な電話からモバイルインターネットアクセスサービスパケット網を通じて、モバイルインターネットアクセスサービスゲートウェイサーバに接続すると、前記自動通訳サーバが、ユーザが通訳可能な言語種、ユーザが通訳可能な文例場面、ユーザが入力可能な文例を該ユーザの電話の表示画面に表示するようにする。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005108]

| | |
|----------|--------------------|
| 1. 変更年月日 | 1990年 8月31日 |
| [変更理由] | 新規登録 |
| 住 所 | 東京都千代田区神田駿河台4丁目6番地 |
| 氏 名 | 株式会社日立製作所 |